

**AUTOMATIC IMAGE CROPPING SYSTEM AND METHOD FOR USE WITH
PORTABLE DEVICES EQUIPPED WITH DIGITAL CAMERAS
CROSS-REFERENCE TO RELATED APPLICATIONS**

[0001] This application claims the benefit of U.S. Provisional Application No. 60/493,232, filed on August 7, 2003. The disclosure of the above application is incorporated herein by reference in its entirety for any purpose.

FIELD OF THE INVENTION

[0002] The present invention generally relates to image processing systems and methods, and relates in particular to automatic image cropping systems and methods for use with portable devices equipped with digital cameras.

BACKGROUND OF THE INVENTION

[0003] Portable devices equipped with cameras, such as Panasonic mobile phones, have been emerging and becoming popular in the market. The resource, such as memory and storage, and the resolution of the camera lens on these portable devices are usually limited. Therefore, their uses are usually limited to the capturing of human objects for wireless image transfer. As a result, most people use the mobile phone camera just for fun. Thus, the camera on a mobile device has not reached its potential. In addition to continuing improving hardware and equipping devices with more memory and storage, more features are called for in order to increase the use of built-in cameras.

[0004] Built-in cameras on a portable device should be able to capture a variety of information from scenes or objects when a user carries it around. Examples are pictures from magazines, billboards, newsletters, catalogs; contact numbers from business cards; URL/phone number from advertisements, and other information. When capturing such information on a portable device, users often have to compliment the focus or the field of angle of the lens. As a result, users typically capture larger than desired area/blocks in the viewing area. These unnecessary regions occupy a large portion of storage space. They also consume bandwidth, thus slowing down the rendering of images on the device's LCD screen. Accordingly, there is need for a way to prevent users from capturing superfluous information.

SUMMARY OF THE INVENTION

[0005] In accordance with the present invention, an automatic image cropping system is for use with a portable device having an image capture mechanism and a limited resource for storing or transmitting captured information. The system includes a region of interest suggestion engine defining plural image region candidates by performing image segmentation on an image stored in digital form. The engine also determines if an image region candidate is likely to be more or less interesting to a user than another image region candidate. The engine further selects an image region candidate determined as likely to be of most interest to the user. In some embodiments, the engine further possesses a training module to track user interaction with the portable device and adjust future determination of likelihood of user interest accordingly.

[0006] Further areas of applicability of the present invention will become apparent from the detailed description provided hereinafter. It should be understood that the detailed description and specific examples, while indicating the preferred embodiment of the invention, are intended for purposes of illustration only and are not intended to limit the scope of the invention.

20 BRIEF DESCRIPTION OF THE DRAWINGS

[0007] The present invention will become more fully understood from the detailed description and the accompanying drawings, wherein:

[0008] Figure 1 is a flow diagram illustrating a method of operation for use with a portable device having a digital camera according to the present invention;

[0009] Figure 2 is a flow diagram illustrating a method of operation for use with a Region Of Interest (ROI) suggestion engine according to the present invention;

[0010] Figure 3 is a flow diagram illustrating a method of training, based on interactive feedback and accumulation, parameters of a cost function employed to suggest ROIs according to the present invention; and

[0011] Figure 4 is a view illustrating an example of segmentation and ROI selection according to the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0012] The following description of the preferred embodiment(s) is merely exemplary in nature and is in no way intended to limit the invention, its application, or uses.

5 **[0013]** The present invention fulfills the needs of users to conserve memory and bandwidth resources by providing an automatic image-cropping scheme to aid users in selecting areas of interest when capturing. This scheme helps to alleviate the problem with memory or bandwidth involved with transmitting an image using a wireless handset. This scheme also
10 facilitates zooming in on a certain object. Thus, the scheme applies to digital still cameras as well.

15 **[0014]** The core components of automatic image cropping are comprised of ROI (region of interest) suggestion engine and a GUI for user confirmation. The suggested ROI from the suggestion engine will be prompted to the user in an easy-to-use graphical interface. As illustrated in Figure 1, as soon as the "shutter" is depressed, resulting in capture of image 10, suggested area 12 (in a highlighted bounding box) is prompted to the user. The user may choose at 14 to select the suggested area, show a next suggested area, or select the entire image without cropping. Based on the
20 user's selection, the selected region can be saved or transmitted without the rest of the image as at 16. The selected area can also be zoomed in depending on the application, which also results in exclusion of image contents outside the confirmed region.

25 **[0015]** Turning to Figure 2, the ROI suggestion engine performs color transformation at step 18, image segmentation at step 20, and entropy based image region candidate and ROI selection at step 22.

30 **[0016]** In step 18, the captured image in RGB format is transformed into HUV (Hue, Saturation and Intensity) format as discussed in A.K. Jain, "Fundamentals of Digital Image Processing", Prentice Hall. The image segmentation and ROI selection algorithm is performed using this color representation.

[0017] In step 20, the image captured on the LCD screen is segmented based on the texture and color consistency. A fuzzy k-mean clustering method can be employed as discussed in A. M. Bensaid, L.O. Hall,

J.C.Bezdek, L.P.Clark, M.L.silbiger, J.A. Arrington and R.F.Murtagh, "Validity-Guided (Re)Clustering with Applications to Image Segmentation", IEEE Trans. on Fuzzy Systems, Vol. 4, No.2, May, 1996. The features used in the clustering method are derived on the color differences of neighboring pixels *i* and *j* defined as

$$C_{diff}(i, j) = \sqrt{(h(i) - h(j))^2 + (u(i) - u(j))^2 + (v(i) - v(j))^2}$$

where *h(i)*, *u(i)* and *v(i)* are the HUV value of pixel *i* and *h(j)*, *u(j)* and *v(j)* are the HUV value of pixel *j*.

[0018] Vectors calculated from Wavelet transform such as

10 Daubechies 3 can be used to represent texture information as discussed in: Robert Porter and Nishan Canagarajar, A Robust Automatic Clustering Scheme for Image Segmentation using Wavelets, IEEE Transactions on Image Processing, Vol. 5, NO. 4, April 1996; Michael Unser, Texture Classification and Segmentation using Wavelet Transform, IEEE Transactions on Image Processing, VOL. 4, NO. 11, November 1995; and T.Chang and C.C. Jay Kuo, Texture Analysis and Classification with Tree-Structured Wavelet Transform, IEEE Transactions on Image Processing, Vol. 2, No. 4, October 1993.

[0019] In step 22, entropy based image region selection is

20 performed in some embodiments. In a preferred embodiment, an algorithm uses entropy as one of plural criteria to determine if a region is more or less interesting to the user. A region with larger entropy contains more information, and thus may be more likely to be of interest to the user.

[0020] The entropy of an image is defined as

25

$$H = - \sum_i h(i) \log_2 h(i)$$

where *h(i)* $i \in I$ is the histogram of the image.

[0021] The higher the entropy, the richer the colors are, and it is assumed that the region with the highest entropy is likely to be the region of interest to the user. The candidate regions are generated in the order of entropy. Considering that human perception can be different from the pure idea of richness in information measured by entropy, these candidates are selected based on several other criteria. Mainly, the size and location of the

candidate areas relative to the entire viewing area are considered. Consequently, a cost function is defined as

$$C = \frac{\alpha}{H_H + H_U + H_V} + \frac{\beta}{A_{ratio}} + \gamma \left(\frac{|X_c - I_{cx}|}{w} + \frac{|Y_c - I_{cy}|}{h} \right)$$

where H_H, H_U, H_V are the entropy of sub-images H, U and V respectively.

5 $A_{ratio} = \frac{Area_{ROI}}{Area_{image}}$ is the area ratio of the ROI and the whole image. x_c, y_c is the center of the ROI while I_{cx}, I_{cy} is the center of the captured image. w, h are the width and height of the lens viewing area, respectively. α, β, γ are normalizing weights. The region with the lowest cost will be prompted to the user first. Camera sensor data (such as user focus area, camera orientation, lens 10 aperture, etc.) may also be used in the suggestion engine.

[0022] The selection of parameters α, β, γ can be based on the characteristics of the camera and the habits of the user. For example, a camera lens with a macro may be able to capture an interested region in relatively larger scale. Therefore, the weight of A_{ratio} can be slightly higher. In 15 yet another example, if a user always saves the entire captured image, the weight of A_{ratio} will out-weight any other parameters ($\alpha=0, \beta=1, \gamma=0$) (i.e., the automatic cropping is turned off). Therefore, human behaviors and habits can be recorded and used to automatically adjust the parameters through a training process that involves interactive feedback and accumulation. The 20 details are illustrated in Figure 3. Initially, the parameters are set empirically to normalize and balance all three components that contribute to the cost: entropy (E), area ratio (A) and center distance (D). In an interactive feedback process, with each captured image 24, segmented blocks are identified in step 20 and four lists of these blocks are generated at step 26 according to E, 25 A, D and their total cost: $\alpha E + \beta A + \gamma D$. Blocks are suggested based on their costs at step 28. The suggested blocks are available for viewing and selection, with the user selecting and confirming a region of interest at step 30. If the user does not select the first suggested region of interest, the three components E, A, D are analyzed on the selected block at step 32 and 30 parameters are adjusted accordingly at steps 34A-34C. It is envisioned that various embodiments can analyze the components on a block in various

ways. For example, a block rejected by a user can be analyzed to incorporate negative feedback. A block selected by the user after rejection of an automatically selected block can alternatively or additionally be analyzed to incorporate positive feedback. It is also possible that user confirmation of an 5 automatically selected block can result in the automatically selected block being analyzed to incorporate positive feedback. Thus, the method in Figure 3 can be modified and supplemented in various ways as will be readily apparent to one skilled in the art. A picture does not necessarily yield the highest entropy when the image with combination of text and pictures is being 10 processed at grey scale level and the text region is captured out of focus (blurred). Pre-processing (smoothing) can be performed to eliminate noise in blurred text histograms.

[0023] Figure 4 is an example of an image captured using a low-end camera (Sharp) plugged into a Sharp Zaurus PDA. The segmentation result 15 is overlaid in the figure. Using the cost function defined above, the area of the picture in the image is selected first as the region of interest, as illustrated with bounding box 12, which has a different display property than bounding boxes 36A-36G used to simultaneously identify other image region candidates. In other words, the automatic image cropping engine shows that the picture area 20 is more likely to be the image region of interest to the user. Consequent actions can be taken upon the user's confirmation: save area, transmit this area (on a mobile phone), or zoom in this area.

[0024] It is envisioned that the user can shift focus between identified regions, and that the region having the focus will have a display 25 property making it distinguishable from other image region candidates. Ranking the regions by entropy or lowest cost facilitates focus shifting by allowing the user to navigate from region to region with few or simplified physical interface components. In some embodiments, bounding boxes are used to indicate the image region candidates, with the hue of a bounding box 30 around an image region candidate that has the focus being different from a hue of bounding boxes about image region candidates that do not have the focus. Example hues are red and green, but it is envisioned that other hues may be used, and that users, such as red-green color blind users, may be given the ability to select to use different display properties. For example,

users may be permitted to select that bounding boxes or other indicators have a relatively more bold appearance when receiving the focus, or that such indicators exhibit different visual patterns. Additional or alternative display properties can also be used. For example, the entire image may be 5 presented as a thumbnail, with the currently selected image region candidate primarily displayed in the active display. Also, indicators, such as bounding boxes, blocks, or lines, may be provided to the thumbnail to show image region candidates with differing display properties. Further, image contents outside all image region candidates may be permitted to blink, while image 10 region candidates not having the focus are steadily rendered in black and white, and the currently selected candidate region is steadily rendered in color. Yet further, the active display of the device GUI may simply display one image region candidate at a time, with the entire image being treated as one of the image region candidates. Further still, the portable device may provide 15 mechanisms (e.g., cursor, arrow button, jog dial, etc.) for users to browse through and select candidate regions. Moreover, various alternative and additional ways to accommodate user browsing, navigation, and selection of image region candidates are envisioned as will be readily apparent to one skilled in the art.

20 [0025] The automatic image cropping scheme of the present invention can be used in a low-resource camera device, such as mobile phone or PDA equipped with a camera, to identify regions of interest from a captured image, and only save a user desired region/block in order to save memory resource on the device.

25 [0026] The algorithm designed for color images and the ROI suggestion engine based on entropy therefore provides intelligence that is closer to a human's perception when capturing an object in the viewing area. Yet, the algorithm is simple to implement with less computational intensity on a low resource device.

30 [0027] The description of the invention is merely exemplary in nature and, thus, variations that do not depart from the gist of the invention are intended to be within the scope of the invention. Such variations are not to be regarded as a departure from the spirit and scope of the invention.